The Identification of MicroRNAs in a Genomically Unstable Region of Human Chromosome 8q24

Konrad Huppi,[1] Natalia Volfovsky,[3] Timothy Runfola,[1] Tamara L. Jones,[1] Mark Mackiewicz,[1] Scott E. Martin,[1] J. Frederic Mushinski,[2] Robert Stephens,[3] and Natasha J. Caplen[1]

[1]Gene Silencing Section, Genetics Branch and [2]Laboratory of Cancer Biology and Genetics, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland; and [3]Advanced Biomedical Computing Center, National Cancer Institute-Frederick/Science Applications International Corporation-Frederick, Inc., NIH, Frederick, Maryland

Abstract

The *PVT1* locus is identified as a cluster of T(2;8) and T(8;22) ''variant'' *MYC*-activating chromosomal translocation breakpoints extending 400 kb downstream of *MYC* in a subset (≈20%) of Burkitt's lymphoma (vBL). Recent reports that microRNAs (miRNA) may be associated with fragile sites and cancer-associated genomic regions prompted us to investigate whether the *PVT1* region on chromosome 8q24 may contain miRNAs. Computational analysis of the genomic sequence covering the *PVT1* locus and experimental verification identified seven miRNAs. One miRNA, *hsa-miR-1204*, resides within a previously described *PVT1* exon (1b) that is often fused to the immunoglobulin light chain constant region in vBLs and is present in high copy number in *MYC/PVT1*–amplified tumors. Like its human counterpart, mouse *mmu-miR-1204* represents the closest miRNA to *Myc* (~50 kb) and is found only 1 to 2 kb downstream of a cluster of retroviral integration sites. Another miRNA, *mmu-miR-1206*, is close to a cluster of variant translocation breakpoints associated with mouse plasmacytoma and exon 1 of mouse *Pvt1*. Virtually all the miRNA precursor transcripts are expressed at higher levels in late-stage B cells (including plasmacytoma and vBL cell lines) compared with immature B cells, suggesting possible roles in lymphoid development and/or lymphoma. In addition, lentiviral vector-mediated overexpression of the *miR-1204* precursor (human and mouse) in a mouse pre–B-cell line increased expression of *Myc*. High levels of expression of the *hsa-miR-1204* precursor is also seen in several epithelial cancer cell lines with *MYC/PVT1* coamplification, suggesting a potentially broad role for these miRNAs in tumorigenesis.

(Mol Cancer Res 2008;6(2):212–21)

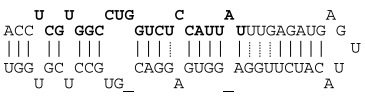
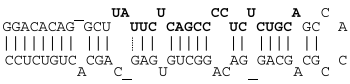
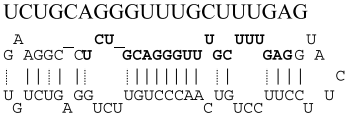
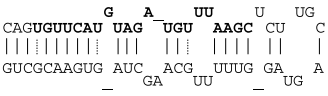
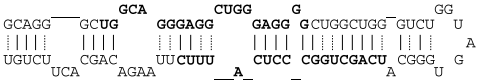
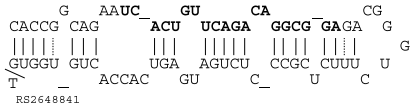
Introduction

For years, noncoding RNAs have represented an exception to the ''Central Dogma'' theory that RNA predominately encodes protein. Recently, a new series of noncoding RNAs have been discovered in a broad range of species, termed primary microRNAs (miRNA) that are processed into mature miRNAs of ~21 to 22 nucleotides (nt) via an intermediate precursor miRNA (1). Analysis of the distribution of miRNA genes in the human genome has suggested that many are found within genomically unstable regions. This finding has strengthened the argument that deregulated expression of miRNAs and/or their targets could be critical components in many disease processes, including cancer (2). For example, a cluster of miRNAs resides on human chromosome 13q14, a frequently deleted chromosomal region in cancer (3). The absence of involvement of any protein-encoding gene in this region coupled with the location of *hsa-miR-15* and *hsa-miR-16* within a 30-kb–deleted segment in chronic lymphocytic leukemia has suggested that, at least in this context, these miRNAs can function as tumor suppressors (4). Other miRNAs that have been implicated in cancer include *hsa-let-7, hsa-mir-155*, and a cluster of miRNAs found within chromosome 13q31 that are frequently amplified in tumors from patients with various forms of lymphoma (5-9).

An early example of a connection between cancer and chromosomal breakpoints are translocations associated with human chromosome 8q24 containing the gene encoding the proto-oncogene transcription factor *MYC* (Supplementary Fig. S1A; ref. 10). In 80% of Burkitt's lymphoma (BL), a T(8;14) chromosomal translocation juxtaposes *MYC* to the *IGH* locus. Variant translocations [T(2;8) or T(8;22)] are found in the remaining 20% of BLs, with breakpoints extending 400 kb downstream of *MYC* in a region referred to as *PVT1*. Because similar clusters of *PVT1*-associated chromosomal breakpoints or retroviral integration sites are found in mouse plasmacytomas (PCT) and rat immunocytomas, it has been suggested that altered expression of *MYC* is the target of this genomic instability (11). For example, *PVT1*-involved translocations are often accompanied by a change in the initiation of *MYC* transcription from the normal P2 promoter to the stronger P1

Received 3/1/07; revised 10/4/07; accepted 10/22/07.
Grant support: Intramural Research Program [Center for Cancer Research, National Cancer Institute (NCI)] of the NIH. Contract N01-CO-012400 supported work conducted by N. Volfovsky and R. Stephens, Advanced Biomedical Computing Center NCI-Frederick/Science Applications International Corporation-Frederick, Inc., NCI, NIH.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
Note: Supplementary data for this article are available at Molecular Cancer Research Online (http://mcr.aacrjournals.org/).
Requests for reprints: Konrad Huppi, Gene Silencing Section, Genetics Branch, Center for Cancer Research, National Cancer Institute, NIH, Building 37, Room 6128, Bethesda, MD 20892. Phone: 301-402-7506; Fax: 301-402-3241. E-mail: huppi@helix.nih.gov
Copyright © 2008 American Association for Cancer Research.
doi:10.1158/1541-7786.MCR-07-0105

Table 1. Experimentally Predicted and Validated miRNAs within the *PVT1* Region on Human chromosome 8q24

miRNA annotation Coordinates chromosome 8	Predicted sequence 5′-3′ Predicted stem loop 5′-3′	Energy	Known related seeds
hsa-miR-1204 128877393- 128877413	UCGUGGCCUGGUCUCCAUUAU 	−26.8	AUUCU mmu-miR-698
hsa-miR-2PVT1 129010218- 129010198	UAUUCUCAGCCCCUCUCUGCA 	−37	
Reverse strand *hsa-miR-1205* 129042068- 129042087	UCUGCAGGGUUUGCUUUGAG 	−23.8	CUGCAG hsa-miR-17-3p
hsa-miR-1206 129090329- 129090349	UGUUCAUGUAGAUGUUUAAGC 	−19.4	
hsa-miR-1207-5p *hsa-miR-1207-3p* 129130587-129130607 129130631-129130648	UGGCAGGGAGGCUGGGAGGGG UCAGCUGGCCCUCAUUUC 	−38.8	UCAGCU rno-miR-337 CAGCUG mmu-miR-763 GCUGGC hsa-miR-565
hsa-miR-1208 129231543-129231616	UCACUGUUCAGACAGGCGGA 	−25.7	

promoter, which may lead to an overall increase in *MYC* expression (reviewed in ref. 10). However, evidence that additional factors are involved in *MYC* deregulation also exists, including the observation that *MYC* transcription levels vary immensely from one tumor to another, independent of breakpoint location (12).

A plausible explanation for the interaction between *MYC* and *PVT1* came with the discovery of *PVT1*-associated transcripts (13, 14) that initiate ~40 to 60 kb downstream of *MYC* in the same transcriptional orientation. Despite the presence of alternatively spliced transcripts, variant chromosomal translocations [T(8;22) or T(2;8)] that juxtapose *IG* λ or κ light chain genes with *PVT1* result in chimeric transcripts of 0.9 to 1.2 kb, which contain *PVT1* exons 1a or 1b spliced directly to the constant region of IG λ or κ (Supplementary Fig. S1A; refs. 15, 16). Although these fusion transcripts are relatively abundant (as a consequence of the proximity to the *IG* enhancer), no protein product(s) have been found and therefore no biological role has been identified. In a study of multiple myeloma, 16% (7 of 42) of patients had rearrangements of the *PVT1* region (independent of *IG* translocation) and this correlated with refractoriness to therapy (17). In variant translocations of mouse PCTs, similar small chimeric transcripts are formed, which contain exon 1a fused directly to the IG light chain constant region. However, as seen with human *PVT1*, no protein product has been identified in mouse for either the fusion products or the normal *PVT1* transcript. Although both human and mouse *PVT1* are transcribed in the same direction as *MYC*, and in virtually all tissues, differences in the location (with reference to *MYC*) and the number of exons suggest that these transcriptional units may not be the same.

Genomic alterations of the chromosome 8q24 region have also been implicated in the development of solid tumors. Coamplification of human *MYC* and *PVT1* seem to correlate with rapidly growing and progressive breast cancer and, as a prognostic factor, this amplicon has been associated with poor clinical survival in postmenopausal or HER2-positive breast cancer patients (18). In breast cancer cell lines with *MYC/PVT1* amplification, usually the entire *MYC* and *PVT1* regions are coamplified, although microdeletions in the region of *PVT1*

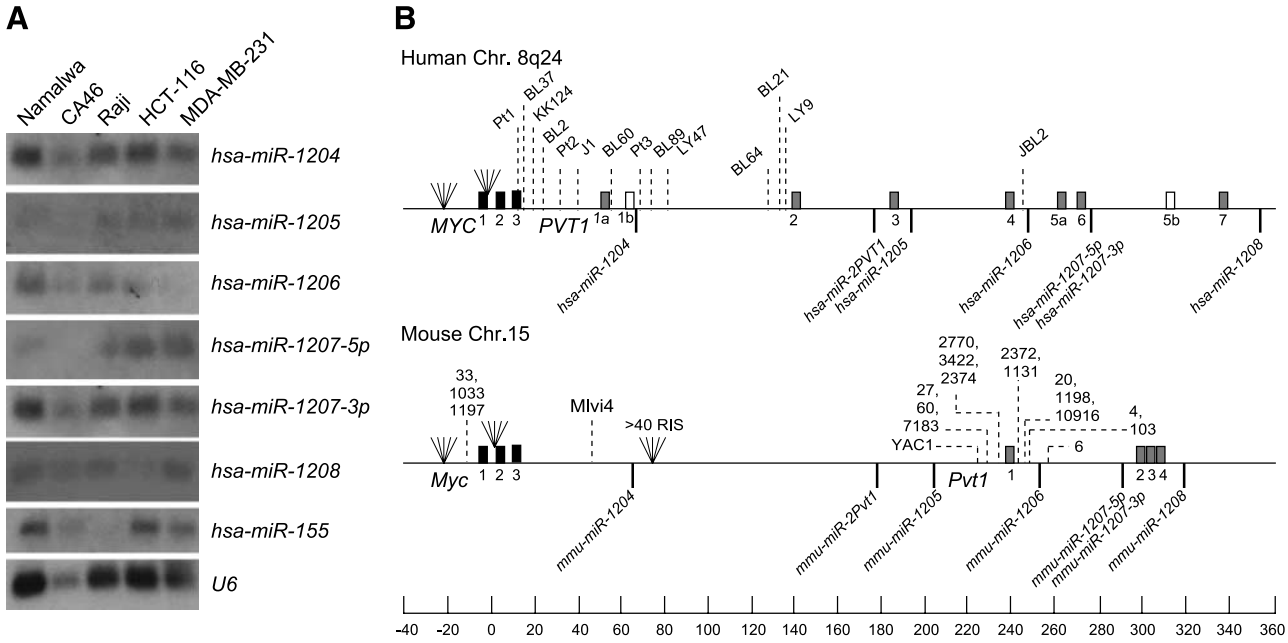


FIGURE 1. The location and validation of miRNAs within chromosome 8q24 region. **A.** Northern blot validation of the computationally predicted miRNAs from the chromosome 8q24 region. U6 was used as a control for loading and expression levels were compared with *hsa-miR-155*. Note that CA46 RNA is underloaded and Namalwa RNA is overloaded. Optimal exposures are shown and exposure times varied from 24 h to 1 wk. **B.** Diagrammatic representation of the position of validated miRNAs associated with the *PVT1* regions of human and mouse. The alignments of precisely mapped BL and PCT breakpoints or retroviral integration sites (i.e., Mlvi4; ref. 11) are shown above the chromosome line. Instances of multiple retroviral integration sites or breakpoints beyond the resolution of the figure are denoted by the symbol with five lines.

have also been seen in some tumor cell lines (i.e., COLO320-HSR and SKBR3; refs. 19, 20; Supplementary Fig. S1B).[4] Nevertheless, the absence of uniformly deregulated *MYC* expression in many different types of epithelial and hematopoietic cancers (ovarian, colorectal, lung, head and neck, leukemia, and lymphoma) exhibiting coamplified *MYC/PVT1* suggests a complex role for this region in tumorigenesis (20-26).

MicroRNAs exert their effect on gene expression through a posttranscriptional gene silencing mechanism termed RNA interference where gene expression is down-regulated through either cleavage of the target transcript or, more typically, through suppression of protein translation. The ≈22-nt mature miRNAs mediate their effect on gene expression through the formation of ribonucleotide-protein complexes that use the miRNA as a guide for the alignment of the complex with a target transcript (27). The sequence alignment between a miRNA and a mRNA transcript is a critical feature of this interaction. To date, little is understood about this alignment and its functional consequence, but minimal requirements seem to include an exact sequence alignment of ≈7 nt (at the 5′ end of the mature miRNA) known as the seed sequence, and a variable number of matches and mismatches between the sequence target region and the miRNA for the remaining ≈15 nt. An important consequence of this imperfect sequence alignment is that each miRNA has the potential to interact with many target transcripts. Thus, a single miRNA could influence the expression of multiple critical genes in a transformation pathway, for example, by directly or indirectly affecting *MYC* expression. In this study, we have tested our hypothesis that the *PVT1* region harbors previously unidentified miRNAs based on the following observations, (*a*) no *PVT1* protein product has been found that might interact with *MYC*, (*b*) the observed differences in location of the mouse and human *PVT1* transcriptional units lessen the importance of a conserved *PVT1* transcript, and (*c*) clusters of miRNAs are frequently associated with regions of genomic instability (2, 28).

Results

In silico *Identification of PVT-Associated miRNAs*

One explanation for the role of the *PVT1* region in human vBL and mouse PCTs as well as epithelial cancers is the possibility that the *PVT1* noncoding RNA may simply be a by-product of a transcriptionally active region harboring other transcriptional units, including miRNAs. In addition to our own cloning studies (Supplementary Fig. S1A; refs 14, 24), a survey of human and mouse *PVT1*-associated sequences[5] shows extensive alternative splicing of transcripts from this region. To test the hypothesis that *PVT1* contains heretofore uncharacterized miRNAs, we analyzed the 400-kb genomic sequence 3′ of MYC, which includes the entire human *PVT1* locus (University of California Santa Cruz,

[4] P. Meltzer, personal communication.

[5] http://moult.umbi.umd.edu/human2004 and http://www.gsc.riken.go.jp/e/FANTOM

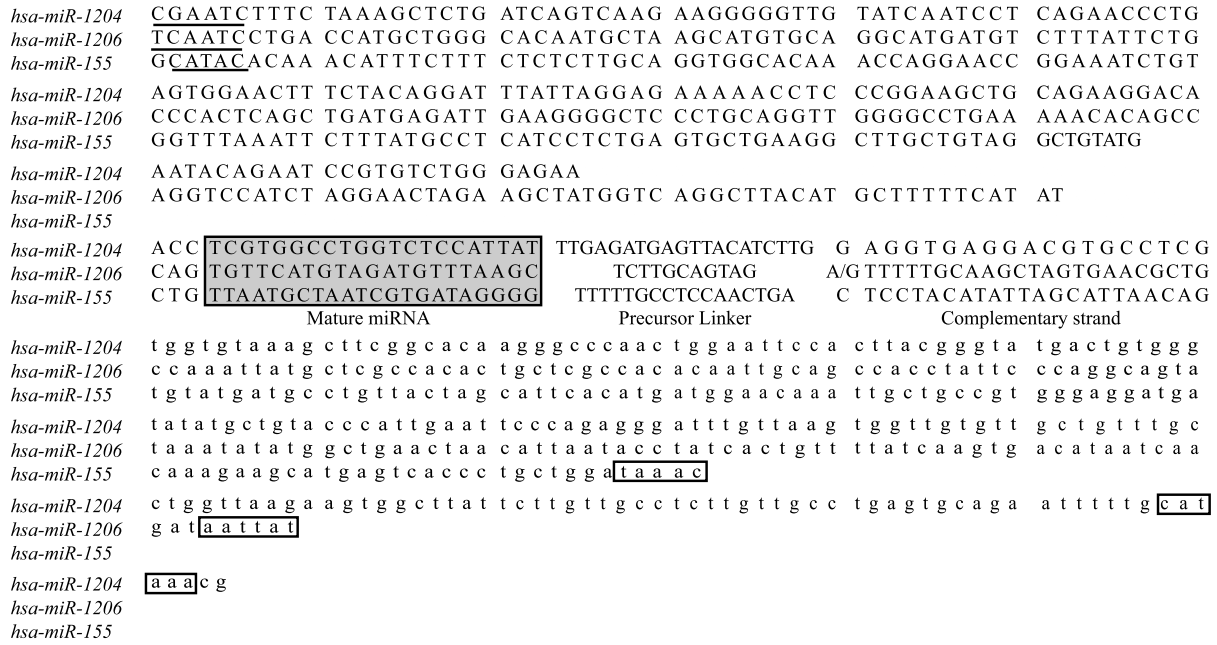


FIGURE 2. Precursor sequences for *hsa-miR-1204* and *hsa-miR-1206* compared with *hsa-miR-155*. 5′ RACE analysis was conducted to obtain suitable PCR products for sequencing. Gray-boxed area, sequences of *hsa-miR-1204, hsa-miR-1206*, and *hsa-miR-155*. A sequence variation (A/G) is found at the beginning of the complementary region for *hsa-miR-1206*. Potential initiation start sites as defined by consensus sequences are underlined. Sequences in the 3′ direction beyond the complimentary strand are predicted, and therefore, are indicated by lower case with potential polyadenylate signals (*boxed*) denoted.

hg17, chr8: 128,830,208-129,253,652) and in multiple species by computational procedures, similar to previously described approaches (29-32). Principally, sequence conservation (primarily putative seed sequences) and secondary hairpin-facilitating structural properties were used as the main criteria for the selection of putative miRNA precursor structures. Conserved segments of 60 bp or more from the *PVT1* region were extracted from the University of California Santa Cruz PhastCons annotation[6] based on the multiple alignments of six species (human, chimpanzee, dog, mouse, rat, and chicken). Any PhastCons elements separated by the gap of <20 bp were merged into one segment for structural analysis. A sliding window of 80 bp with an increment of 10 bp was scanned along the conserved segments. The sequence windows were folded using the RNAfold program (33) and the stem-loop structures with a total length of >50 bp and with at least 65% of paired bases were selected. The remaining sequences were refolded and the structures with a folding free energy of over a maximum of −15 kcal/mol were selected for experimental validation. The calculation of the free energy of the predicted secondary precursor structures was coupled with the ratio of the paired bases to the total length of the sequence. This analysis predicted 13 potential mature miRNA sequences (Table 1; Supplementary Fig. S2).

Validation of PVT1 *miRNAs by Northern Analysis*

Northern blot hybridization was used to test for expression of each of the miRNAs identified by the bioinformatics process using size-selected RNA (less than ≈30 nt) from BL, breast, and colon cancer cell lines with radioactively labeled ribonucleotide probes corresponding to each of the computationally predicted mature miRNA sequences (Supplementary Methods). The expression of six mature miRNA sequences was confirmed (Fig. 1A). For the remaining seven miRNAs that were apparently not expressed, probes were made in the antisense direction and rechecked by Northern blot analysis to eliminate strand bias. Only *hsa-miR-2PVT1* was expressed in an antisense direction and at very low levels (detected only after 1 month of exposure; data not shown). Overall, 7 of the 13 computationally predicted mature miRNA sequences were validated experimentally. Some miRNAs were expressed abundantly (*hsa-miR-1204, hsa-miR-1207-5p*, and *hsa-miR-1207-3p*) with similar patterns of expression among the cell lines, whereas others seem to be expressed at lower levels (*hsa-miR-1205* and *hsa-miR-1208*). We also observed some examples of increased or decreased expression on an individual basis within the cell lines used. For example, *hsa-miR-1207-3p* is expressed abundantly in the colon cancer cell line HCT-116. Conversely, *hsa-miR-1206* seems to be expressed only at low levels in the MDA-MB-231 breast cancer cell line. To establish a basis for comparison, the level of expression of a known human miRNA, *hsa-miR-155*, was also included as part of the Northern blot analysis.

We compared the sequences of these new mature miRNAs with previously reported miRNA sequences (MiRBase, Release 9.0).[7] Examination of the seed sequences for the validated *PVT1*-associated miRNAs revealed only minimal overlap (±1) with seed sequences from already annotated miRNAs (Table 1).

[6] http://www.genome.ucsc.edu

[7] http://microrna.sanger.ac.uk/

However, the seed sequence for *hsa-miR-1205* is identical to the seed sequence for *hsa-miR-17-3p*, which is expressed in HeLa cells (34). A sequence within *hsa-miR-2PVT1* is the same as the seed sequence of the miRNA *mmu-miR-698*, a recently cloned miRNA from mouse embryos (35). The *hsa-miR-1207-3p* sequence is related to a group of mammalian miRNAs with overlapping seeds [*rno-miR-337* (28), *mmu-miR-763* (36), and *hsa-miR-565* (37)]. The lengths of the predicted precursor sequences are between 58 and 72 nt for most of the miRNAs. However, two miRNAs *hsa-miR-1207-5p* and *hsa-miR-1207-3p*, which we originally considered to be independent overlapping precursors, may reside on complementary strands of a single, longer (87 nt) hairpin structure (Table 1). As mentioned above, both these miRNAs are expressed similarly across the cell lines when examined by Northern blot analysis (Fig. 1A), suggesting a common promoter and transcriptional regulatory unit. Although such a precursor structure is relatively rare, previous reports have identified other examples of complementary overlapping miRNAs (e.g., *let-7d* and *let-7d**; ref. 38). How these miRNA species are processed and used is probably complex and will require further investigation.

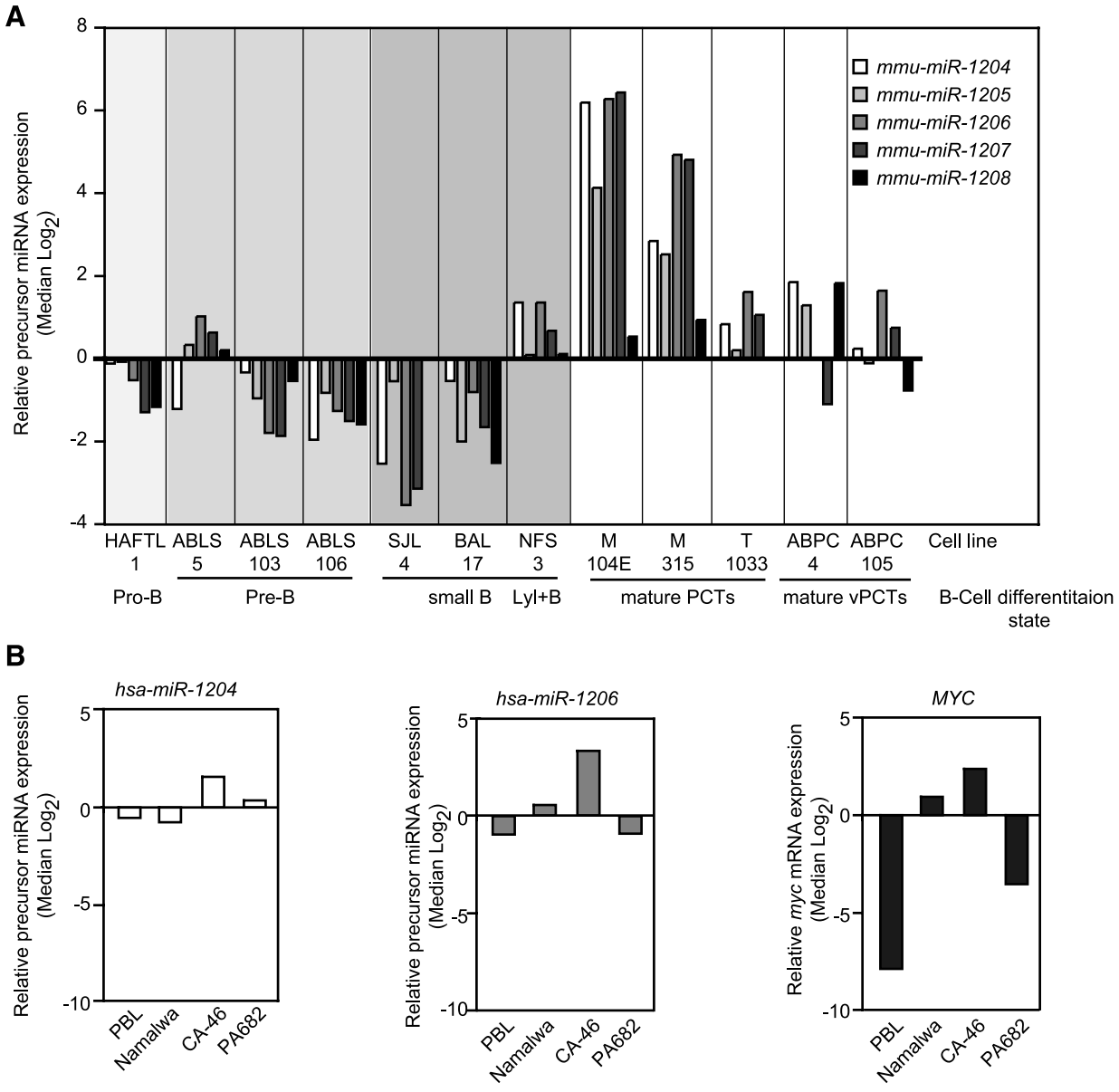


FIGURE 3. Quantitative RT-PCR analysis of *miR-1204* and *miR-1206* precursor expression in human and mouse B-cell lines and lymphoid tumors (**A**). The expression of *mmu-miR-1204, mmu-miR-1206*, and *MYC* in mouse B-cell lines; pro-B (HAFTL1), pre-B (ABLS5, ABLS103, ABLS105), small B (SJL4, BAL17), Ly1+ B (NFS3), mature PCTs (M104E M315, T1033), and mature vPCTs (ABPC4, ABPC105; ref. 46). **B.** The expression of *hsa-miR-1204, hsa-miR-1206*, and *MYC* in the human BL lines; Namalwa, CA46, and the vBL line PA682 compared with normal peripheral blood lymphocytes. Expression levels were normalized to actin and the log 2 expression levels were determined from the ratio of the sample value to the normal tissue value (cDNA synthesis and PCR amplification were repeated three to four times).

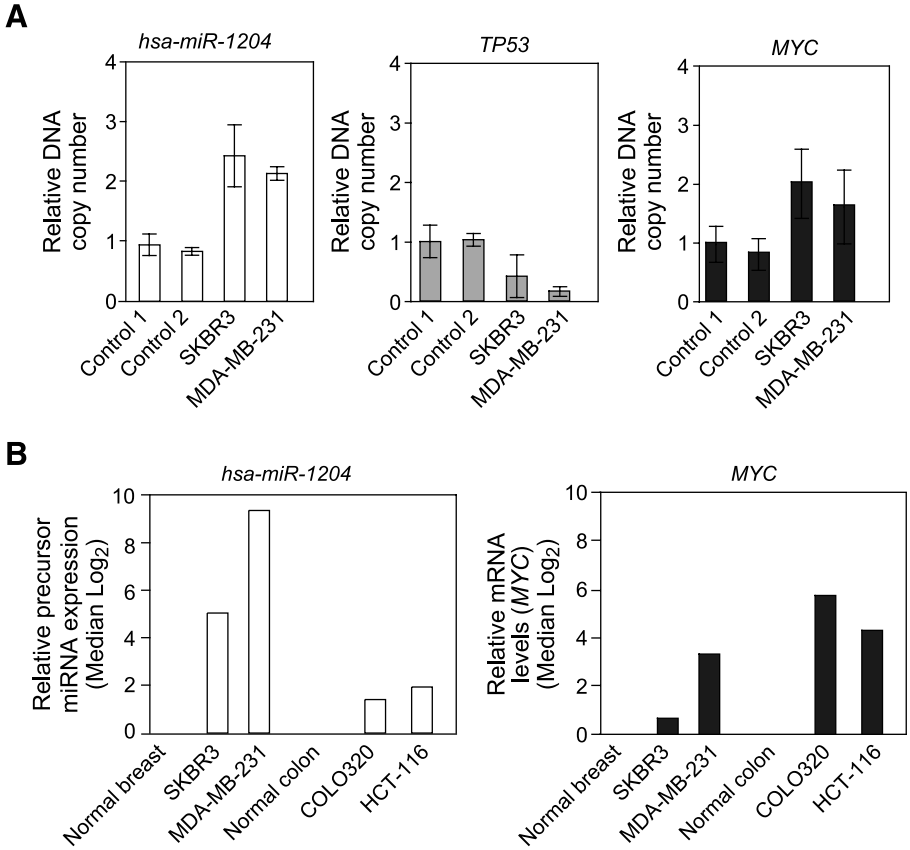


FIGURE 4. PCR-based analysis of DNA copy number and quantitative RT-PCR transcription of *MYC* and *hsa-miR-1204* in *MYC/PVT1* in amplified tumor cell lines or BL cell lines. **A.** DNA copy number for *MYC* and *hsa-miR-1204* in two breast cancer cell lines, SKBR3 and MDA-MB-231, compared with normal individuals [(control 1) 603 (control 2) 610]. DNA copy number for *TP53* was used as a control. **B.** Expression of precursor *hsa-miR-1204* in two breast cancer cell lines (MDA-MB-231 and SKBR3) and in two colon cell lines (HCT-116 and COLO-320) compared with normal breast or colon tissue. Expression levels were normalized to actin and the log 2 expression levels were determined from the ratio of the sample value to the normal tissue value. Expression levels were normalized to actin and the log 2 expression levels were determined from the ratio of the sample value to the normal tissue value (cDNA synthesis and PCR amplification were repeated three to six times).

PVT1-Associated miRNAs and the MYC/PVT1 Genomic Locus

The location of the validated miRNAs was positioned with respect to the human and mouse *PVT1* loci and to known breakpoints associated with human BLs, mouse PCTs, and retroviral integration site (Fig. 1B). Although there are significant differences in the positions of the annotated exons of the human and mouse *PVT1* transcripts, the relative positions of the miRNAs are highly conserved between species. Interestingly, one of the identified miRNA sequences, *hsa-miR-1204*, is located 60 kb downstream of *MYC* and very close to a small CpG island. The mouse counterpart, *mmu-miR-1204*, is located in the same region and close to a large cluster (>40) of retroviral integration site identified by GENETRAP technology[8] and close to a small cluster of T(15;16) breakpoints in vPCTs (39). Furthermore, a sequence-based BLAST search confirmed that the *hsa-miR-1204* sequence overlaps precisely with the *PVT1* exon 1b. As mentioned previously, this exon is found frequently as part of the *PVT1/IG* fusion transcripts generated by vBL breakpoints (Fig. 1B; Supplementary Fig. S1A), and a reverse transcription-PCR assay using primers from the *PVT1* exon 1b and *IG* constant regions usually identifies vBL and vPCT breakpoints (15, 16). The location of *hsa-miR-1204* within exon 1b of *PVT1* is of particular interest, as the predicted *hsa-miR-1204* precursor sequence extends beyond the consensus splice site of the *PVT1* exon 1b. Thus, it is unlikely that *hsa-miR-1204* is merely captured as part of the *PVT1/IG* fusion transcript observed in vBLs. No other miRNAs were associated with annotated *PVT1* exons, although both mouse *miR-1206* and *miR-1205* are found close to the largest cluster of vPCT breakpoints and exon 1 of the mouse *PVT1* cDNA.

Analysis of the hsa-miR-1204 Putative Precursor Sequences

All but one of the identified *PVT1*- associated miRNAs seem to be transcribed in the same orientation indicating shared regulatory elements. To begin to dissect the transcriptional properties of *PVT1*-associated miRNA transcripts, we initially focused our efforts on the precursor transcript associated with *hsa-miR-1204*, as its location at the amino-terminal region of *PVT1* transcription suggests that suitable promoters and transcription initiation signals are likely to be nearby. Initially, we used primer extension from both sense- and antisense-based primers corresponding to *hsa-miR-1204* to confirm the presence of a precursor miRNA transcript. Transcripts of ~75 nt corresponding to *hsa-miR-1204* were found to be abundant in the antisense direction, whereas virtually no transcripts were found using a sense oligonucleotide (data not shown). Next, we used 5′ rapid amplification of cDNA ends (RACE) analysis to define the start of *hsa-miR-1204*. For comparison, we also synthesized RACE products that would define the transcriptional start sites for *hsa-miR-1206*

[8] http://www.genetrap.org

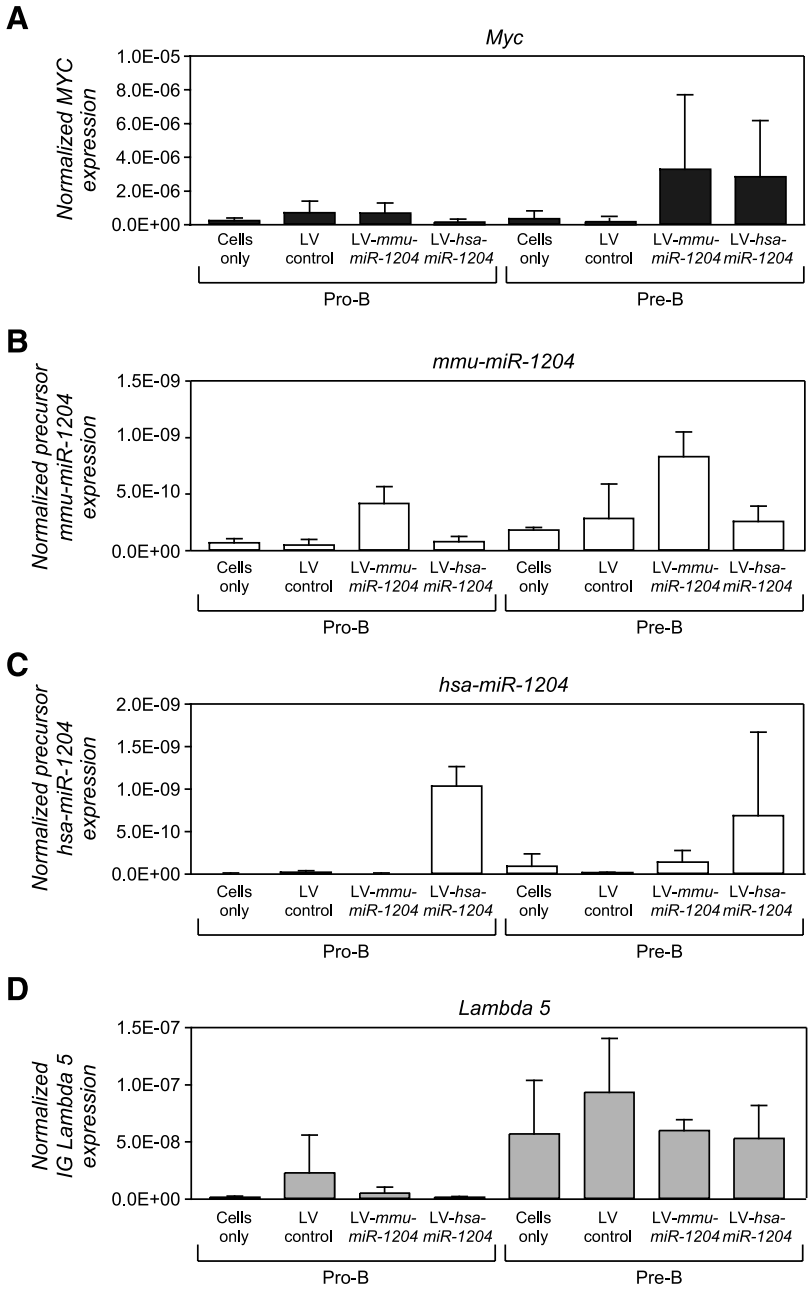


FIGURE 5. Quantitative RT-PCR analysis of *MYC* expression in mouse pro-B or pre-B cells transduced with lentivirus. Pro-B and pre-B cell lines were transduced with a lentiviral vector expressing, from a cytomegalovirus promoter, either a control sequence, the precursor *mmu-miR-1204*, or the precursor *hsa-miR-1204*. Following blasticidin selection of transfectants, quantitative RT-PCR analysis was done for mouse *Myc* (**A**), the precursor *mmu-miR-1204* (**B**), and the precursor *hsa-miR-1204* (**C**). **D.** *IgLL1* ($\lambda 5$) expression was used to confirm the pre-B lineage. Columns, mean of fluorescent units from at least three to four independent experiments; bars, SD.

and *hsa-miR-155* (Fig. 2). Sequence analysis of the RACE products identified ~150 to 175 nt of sequence in the 5′ flanking region of each mature miRNA. In each case, a reasonable consensus transcriptional start sequence ($YYA_{+1}NT$/AYY) is found close to the beginning of the RACE product (40). A sequence variation (A/G) that lies at the beginning of the complementary strand in *hsa-miR-1206* was identified by direct sequencing of PCR products from the peripheral blood lymphocytes of several healthy individuals (Fig. 2). A sequence variant has also been annotated that falls at the 3′ end of the *hsa-miR-1208* precursor (dbSNP:rs2648841). Although functional studies will be necessary to firmly establish whether these sequence variants influence the efficiency of miR processing, it is noteworthy that sequence variations in the 3′ region of *hsa-miR-16-1* and *hsa-miR-15a* have been thought to affect expression in chronic lymphocytic leukemia patients (41). The putative precursor structure for each of these miRNAs was also examined in the mouse, and their conservation of sequence and size strengthened the assignment of precursors to these structures (Supplementary Fig. S3).

Expression Analyses of PVT1*-Associated miRNAs in Normal and Tumor Cells by Quantitative RT-PCR*

There is increasing evidence that many miRNAs play an important role in development and/or differentiation and that their deregulated expression may be a contributing factor

in a variety of cancers. Given the association of the *PVT1* region with BLs and PCTs, we first examined the expression of *mmu-miR-1204* and *mmu-miR-1206* precursors in a series of murine cell lines representing different stages of B-cell differentiation (Fig. 3A; ref. 42). To perform this analysis, we designed and implemented real-time RT-PCR assays using the sequences obtained from the RACE experiments with primers that would recognize the precursor miRNA transcripts in either human or mouse, accordingly. The rationale behind targeting precursors in the real-time assay was to increase the specificity and eliminate size bias in the targeted RNAs. From this analysis, it is apparent that expression of both *mmu-miR-1204* and *mmu-miR-1206* transcripts seems to be higher in mature as opposed to immature stages of B cells. The highest levels, in fact, are found in cell lines expressing IG λ (NFS3, M315 and M104E). In one vPCT cell line (ABPC4), high levels of *mmu-miR-1204* are found but low levels of *mmu-miR-1206* are observed. In contrast, another vPCT cell line (ABPC20) shows high levels of *mmu-miR-1206* and low levels of *mmu-miR-1204*. Although limited samples were examined, this analysis showed no direct correlation between miR-*PVT1*–based miRNA expression and MYC transcription (MYC protein was not assayed). We also quantified the expression of the miRNA precursors corresponding to *hsa-miR-1204* and *hsa-miR-1206* in a limited set of human B-cell tumor samples representing conventional T(8;14) (Namalwa or CA46) and vBL T(8:22; PA682) samples (Fig. 3B). In concert with the findings in the mouse B-cell lines, *hsa-miR-1204* seems to be expressed abundantly in CA46, which produces both IG λ and κ. In contrast, Namalwa, which is also a dual IG λ/κ producer, does not show high levels of *hsa-miR-1204* by quantitative RT-PCR, although *hsa-miR-1206* levels are high (it should be noted that the Northern blot analysis in Fig. 1A seems to show higher levels of *hsa-miR-1204* in Namalwa due to increased RNA loading).

In addition to genomic rearrangements affecting the *PVT1* region, a large number of different types of tumors, both hematopoietic and epithelial, show genomic coamplification of the *MYC/PVT1* region (Supplementary Fig. S1B; refs. 18-26). For example, a recent study of head and neck squamous cell carcinoma found evidence of amplification of *MYC* (38%) or *PVT1* (20%) or both (14%), and concluded that amplification of the *PVT1* region must activate *MYC* as a result of *cis*-acting interaction (22). A similar finding of *MYC* and *PVT1* coamplification (two to five copies) in 8% of primary breast tumors ($n = 311$) showed a significant correlation with rapidly growing and progressive cancer (18). In some breast cell lines (i.e., SKBR3, MDA-MB-231, and MCF-7), detailed comparative genomic hybridization studies have revealed small nonamplified regions (and/or microdeletions) within the *MYC/PVT1* amplicon, suggesting that certain regions may not be coamplified (Supplementary Fig. S1B; ref. 19).[9] Therefore, we examined the copy number of *hsa-miR-1204* compared with *MYC* (Fig. 4). Application of real-time PCR to genomic DNA purified from two breast cancer cell lines (SKBR3, MDA-MB-231) compared with DNA from normal healthy individuals was conducted using primers corresponding to *MYC, hsa-miR-1204*, and *TP53* (as a control). DNA from both SKBR3 and MDA-MB-231 cells displayed a 1.5- to 2-fold amplification of copy number of *MYC* and a 2.0- to 2.5-fold increase in copy number for *hsa-miR-1204*. We also tested *MYC/PVT1*–amplified cell lines compared with normal breast and colon RNA using a quantitative RT-PCR precursor assay for *hsa-miR-1204* and *MYC* transcription (Fig. 4). Enhanced expression of both *MYC* and *hsa-miR-1204* was seen in two colon cell lines (HCT-116, COLO-320) as well as in both of the breast cancer cell lines.

Transduction of Mouse Pre–B-Cell Lines with miR-1204 Results in Overexpression of MYC

We further investigated a possible link between *MYC* and *miR-1204* expression by transducing mouse pro-B and mouse pre-B cells with lentiviral-based vectors expressing either human or mouse *miR-1204* precursors (Fig. 5). Although stably transduced cells showed appropriate expression of the *miR-1204* precursor (Fig. 5B and C), *Myc* expression was enhanced by the overexpression of either the human or mouse *miR-1204* precursor, but only in the pre–B cell (Fig. 5A). This effect on Myc expression is most likely to be indirect as there are no obvious direct binding sites for *miR-1204* within the 3′ untranslated region of *Myc* where one might expect miRNA binding (data not shown).

Discussion

The study of unstable genomic regions has been widely used to identify loci contributing to disease processes, particularly cancer. Genetic rearrangements of human chromosome 8q24 and the syntenic region in mouse (chromosome 15) are associated with lymphoma, and amplification of this same region has also been observed in a variety of epithelial cancers (18-26). The deregulated expression of *MYC*, located within this region, is obviously a key factor in the development of many tumors, but *MYC* does not act alone (reviewed in ref. 43). Thus, many studies have focused on delineating the role of the physically linked *PVT1* locus. Recently, *PVT1* has been found to be one of the most significantly up-regulated transcripts in Eμ-*hsa-miR-155* transgenic mice that readily develop B-cell leukemia/lymphoma (44). The *PVT1* transcript was also found among a small subset of genes that were overexpressed following RNA interference–mediated down-regulation and/or Gleevec inhibition of *BCR/ABL* in chronic myelogenous leukemia (K562) cells (45). In earlier studies, a knockout of the *Pvt1* locus in the mouse was embryonically lethal,[10] suggesting that *PVT1* may play a critical role in normal development. However, the sequence and structural differences in the location of exons between mouse and human *PVT1* transcripts, as well as the vast level of alternative splicing associated with these transcripts, has complicated the interpretation of these results. Furthermore, lack of a protein product for *PVT1* has severely limited the extent of studies, despite its apparent involvement in translocations and amplification in so many different forms of cancer. To begin to answer these questions, we formulated the hypothesis that noncoding miRNAs might be present within the *PVT1* locus and have now validated the presence of seven new

[9] P. Meltzer, personal communication.
[10] K. Huppi, unpublished data.

mature miRNA sequences derived from this region. With the exception of *hsa-miR-2PVT1* (which is expressed at much lower levels than the other miRNAs identified in this study), all of the miRNAs identified are transcribed in the same direction as both *MYC* and *PVT1*.

A number of studies have shown that the normal expression of miRNAs can be affected by chromosomal translocation, amplification, or deletion and that this deregulated expression has the potential to have a direct consequence on the function of these miRNAs (3, 4). Recently, the oncogenic potential of deregulated miRNA expression has also been revealed through the use of a retroviral mutagenesis approach, which shows that retroviral integration of a murine leukemia virus into the *mir-17-92* cluster on mouse chromosome 14 increased expression of the primary miRNA transcript and induced development of T-cell lymphomas (46). The miRNAs identified in this study may also be susceptible to some of the transcriptional alterations previously described for *MYC* and *PVT1* as a result of the chromosomal translocations, amplifications, or retroviral integrations commonly associated with the chromosome 8q24 region (and the syntenic region in mouse). To address the possible effects of previously defined genomic and transcriptional alterations within this region on the miRNAs identified in this study, we focused initially on the expression of *miR-1204* as this miRNA maps closest to *MYC* in human and mouse, partially overlaps with a previously defined exon of human *PVT1* and resides close to a cluster of retroviral integration site in mouse. We observed high levels of the *mmu-miR-1204* precursor in all vBL and vPCT samples. Interestingly, we have observed that the highest levels of *miR-1204* transcription are found in many samples that abundantly express *IG* λ as their light chain (NFS3, M315, M104E, and CA46). If this correlation can be confirmed in additional samples, this may suggest a role *miR-1204* in controlling expression of *IG* λ or *IG* light chain genes in general. We also asked whether genomic and transcriptional changes associated with this region seen in solid tumors could be extended to the *PVT1*-based miRNAs. An increase in copy number and expression for *hsa-miR-1204*, consistent with that seen for *MYC*, was also observed in several breast and colon cell lines. To further substantiate a connection between the *PVT1* based miRNAs and *MYC*, we constitutively overexpressed *miR-1204* (mouse and human) in mouse pro–B-cell or pre–B-cell lines. Increased *Myc* transcription was observed in the pre–B-cell line only, suggesting that a stage-specific down-regulation of a target(s) must have some indirect effect on *Myc* expression. Although further studies will be required to determine the possible function of the miRNAs identified in this study and their role in tumorigenesis, the current study illustrates the value of looking for the presence of regulatory RNAs in chromosomal regions associated with cancer where no protein-encoding gene candidate has formally been identified.

Materials and Methods

Cell Lines

The BL cell lines Raji, Namalwa, PA682, and CA46 were maintained in RPMI 1640 plus 10% fetal bovine serum and are the generous gifts of Drs. G. Tosato, W. Shi, M. Zajac-Kaye, and K. Bhatia [National Cancer Institute (NCI), Bethesda, MD]. The breast cancer cell lines MDA-MB-231, MCF7, SK-BR3, and the colon cancer cell line HCT-116,[11] were maintained in DMEM plus 5% fetal bovine serum. Mouse PCTs and cell lines representing various stages of B-cell development have been described previously (42). Mouse pro-B (BAF3) and pre-B (v-abl-PreB1) cell lines were maintained in RPMI 1640 supplemented with interleukin-3 (10%) for the BAF-3 cells and 2-βME (5mmol/L) for v-abl-PreB1.

Molecular Analysis of miRNAs

Northern analysis was conducted using small RNA (less than ≈30 nt) extracted from total RNA according to the manufacturer's recommendation (Ambion, Inc.) and standard gel electrophoresis, Northern blot transfer, and hybridization to ^{32}P-labeled riboprobes. For additional details, see Supplementary Methods. The methods used for cloning *PVT1* exons, primer extension, and 5′ RACE studies of precursor miRNA sequences and quantitative PCR analysis of miRNA copy number and expression are detailed in Supplementary Methods. The generation of lentiviral-based plasmids expressing *hsa-miR-1204* and *mmu-miR-1204*, viral production, and the selection of stably transduced cell populations is described in Supplementary Methods. The sequences of the oligonucleotide primers used in this study are shown in Supplementary Table S1.

Acknowledgments

We thank Dr. Paul Meltzer (Genetics Branch, Center for Cancer Research, NCI, NIH) for sharing unpublished data; Brady Wahlberg for expert assistance in cloning the human and mouse miRNA constructs; James Owens and Leigh-Ann Cruz of the Laboratory of Cancer Biology and Genetics, CCR, NCI, NIH; and B. Conde and B. Crise (Advanced Technology Program, NCI-Frederick/Science Applications International Corporation-Frederick, Inc.) for expert assistance.

Statement of authorship: K. Huppi, R. Stephens, N. Volfovsky, and N.J. Caplen conceived and designed research. K. Huppi, N. Volfovsky, T. Runfola, T.L. Jones, M. Mackiewicz, and S.E. Martin acquired data. J.F. Mushinski contributed analytic tools. K. Huppi, N. Volfovsky, R. Stephens, and N.J. Caplen analyzed data. K. Huppi, N. Volfovsky, J.F. Mushinski, and N.J. Caplen drafted the manuscript.

[11] http://dtp.nci.nih.gov/

References

1. Du T, Zamore PD. microPrimer: the biogenesis and function of microRNA. Development 2005;132:4645–52.

2. Huppi K, Volfovsky N, Mackiewicz M, et al. MicroRNAs and genomic instability. Semin Cancer Biol 2007;17:65–73.

3. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002;99:15524–9.

4. Cimmino A, Calin GA, Fabbri M, et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 2005;102:13944–9.

5. Johnson SM, Grosshans H, Shingara J, et al. RAS is regulated by the let-7 microRNA family. Cell 2005;120:635–47.

6. Takamizawa J, Konishi H, Yanagisawa K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 2004;64:3753–6.

7. Yanaihara N, Caplen N, Bowman E, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 2006;9:189–98.

8. Volinia S, Calin GA, Liu CG, et al A microRNA expression signature of

human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006; 103:2257–61.

9. He L, Thomson JM, Hemann MT, et al. A microRNA polycistron as a potential human oncogene. Nature 2005;435:828–33.

10. Marcu KB, Bossone SA, Patel AJ. myc function and regulation. Annu Rev Biochem 1992;61:809–60.

11. Lazo PA, Lee JS, Tsichlis PN. Long-distance activation of the Myc protooncogene by provirus insertion in Mlvi-1 or Mlvi-4 in rat T-cell lymphomas. Proc Natl Acad Sci U S A 1990;87:170–3.

12. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med 2006;354: 2419–30.

13. Shtivelman E, Henglein B, Groitl P, Lipp M, Bishop JM. Identification of a human transcription unit affected by the variant chromosomal translocations 2;8 and 8;22 of Burkitt lymphoma. Proc Natl Acad Sci U S A 1989;86:3257–60.

14. Huppi K, Siwarski D, Skurla R, Klinman D, Mushinski JF. Pvt-1 transcripts are found in normal tissues and are altered by reciprocal(6;15) translocations in mouse plasmacytomas. Proc Natl Acad Sci U S A 1990;87:6964–8.

15. Shtivelman E, Bishop JM. Effects of translocations on transcription from PVT. Mol Cell Biol 1990;10:1835–9.

16. Huppi K, Siwarski D. Chimeric transcripts with an open reading frame are generated as a result of translocation to the Pvt-1 region in mouse B-cell tumors. Int J Cancer 1994;59:848–51.

17. Palumbo AP, Boccadoro M, Battaglio S, et al. Human homologue of Moloney leukemia virus integration-4 locus (MLVI-4), located 20 kilobases 3′ of the myc gene, is rearranged in multiple myelomas. Cancer Res 1990;50: 6478–82.

18. Borg A, Baldetorp B, Ferno M, Olsson H, Sigurdsson H. c-myc amplification is an independent prognostic factor in postmenopausal breast cancer. Int J Cancer 1992;51:687–91.

19. Shtivelman E, Bishop JM. The PVT gene frequently amplifies with MYC in tumor cells. Mol Cell Biol 1989;9:1148–54.

20. Feo S, Di Liegro C, Jones T, Read M, Fried M. The DNA region around the c-myc gene and its amplification in human tumour cell lines. Oncogene 1994;9: 955–61.

21. Lancaster JM, Dressman HK, Whitaker RS, et al. Gene expression patterns that characterize advanced stage serous ovarian cancers. J Soc Gynecol Investig 2004;11:51–9.

22. Bhattacharya N, Sabbir MG, Roy A, Dam A, Roychoudhury S, Panda CK. Approximately 580 Kb surrounding the MYC gene is amplified in head and neck squamous cell carcinoma of Indian patients. Pathol Res Pract 2005;201: 691–7.

23. Bakkus MH, Brakel-van Peer KM, Michiels JJ, van 't Veer MB, Benner R. Amplification of the c-myc and the pvt-like region in human multiple myeloma. Oncogene 1990;5:1359–64.

24. Huppi K, Siwarski D, Shaughnessy JD, Jr., Mushinski JF. Co-amplification of c-myc/pvt-1 in immortalized mouse B-lymphocytic cell lines results in a novel pvt-1/AJ-1 transcript. Int J Cancer 1993;53:493–8.

25. Asker C, Mareni C, Coviello D, et al. Amplification of c-myc and pvt-1 homologous sequences in acute nonlymphatic leukemia. Leuk Res 1988;12: 523–7.

26. Storlazzi CT, Fioretos T, Paulsson K, et al. Identification of a commonly amplified 4.3 Mb region with overexpression of C8FW, but not MYC in MYC-containing double minutes in myeloid malignancies. Hum Mol Genet 2004;13: 1479–85.

27. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev 2006;20:515–24.

28. Calin GA, Sevignani C, Dumitru CD, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 2004;101:2999–3004.

29. Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP. Vertebrate microRNA genes. Science 2003;299:1540.

30. Lim LP, Lau NC, Weinstein EG, et al. The microRNAs of *Caenorhabditis elegans*. Genes Dev 2003;17:991–1008.

31. Altuvia Y, Landgraf P, Lithwick G, et al. Clustering and conservation patterns of human microRNAs. Nucleic Acids Res 2005;33:2697–706.

32. Xie X, Lu J, Kulbokas EJ, et al. Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. Nature 2005; 434:338–45.

33. Hofacker IL. Vienna RNA secondary structure server. Nucleic Acids Res 2003;31:3429–31.

34. Lagos-Quintana M, Rauhut R, Meyer J, Borkhardt A, Tuschl T. New microRNAs from mouse and human. RNA 2003;9:175–9.

35. Mineno J, Okamoto S, Ando T, et al. The expression profile of microRNAs in mouse embryos. Nucleic Acids Res 2006;34:1765–71.

36. Berezikov E, van Tetering G, Verheul M, et al. Many novel mammalian microRNA candidates identified by extensive cloning and RAKE analysis. Genome Res 2006;16:1289–98.

37. Cummins JM, He Y, Leary RJ, et al. The colorectal microRNAome. Proc Natl Acad Sci U S A 2006;103:3687–92.

38. Houbaviy HB, Murray MF, Sharp PA. Embryonic stem cell-specific microRNAs. Dev Cell 2003;5:351–8.

39. Axelson H, Panda CK, Silva S, et al. A new variant 15;16 translocation in mouse plasmacytoma leads to the juxtaposition of c-myc and immunoglobulin λ. Oncogene 1991;6:2263–70.

40. Smale ST, Baltimore D. The "initiator" as a transcription control element. Cell 1989;57:103–13.

41. Calin GA, Ferracin M, Cimmino A, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353:1793–801.

42. Mushinski JF, Davidson WF, Morse HC III. Activation of cellular oncogenes in human and mouse leukemia-lymphomas: spontaneous and induced oncogene expression in murine B lymphocytic neoplasms. Cancer Invest 1987;5: 345–68.

43. Adhikary S, Eilers M. Transcriptional regulation and transformation by Myc proteins. Nat Rev Mol Cell Biol 2005;6:635–45.

44. Costinean S, Zanesi N, Pekarsky Y, et al. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(μ)-miR155 transgenic mice. Proc Natl Acad Sci U S A 2006;103:7024–9.

45. Zhelev Z, Bakalova R, Ohba H, et al. Suppression of bcr-abl synthesis by siRNAs or tyrosine kinase activity by Glivec alters different oncogenes, apoptotic/ antiapoptotic genes and cell proliferation factors (microarray study). FEBS Lett 2004;570:195–204.

46. Wang CL, Wang BB, Bartha G, et al. Activation of an oncogenic microRNA cistron by provirus integration. Proc Natl Acad Sci U S A 2006; 103:18680–4.